

## DOCUMENT INFORMATION MANAGEMENT SYSTEM

### FIELD OF THE INVENTION

The present invention relates to a document information  
5 management system enabling search of documents only by specifying  
an arbitrary word as a keyword in a hypertext or a non-hypertext  
document displayed on a screen.

### BACKGROUND OF THE INVENTION

10 Conventionally, a form of a document commonly used in the  
world has generally been a paper document, but progress of the  
Internet or intranet has given a birth to a hypertext as an  
electronic document which can compete for the first time against  
the paper document. This hypertext document is a document  
15 written in HTML in a narrow sense, and includes all the electronic  
documents which can link between documents such as PDF documents  
by Adbe and a portion of an SGML document by CALS (namely, between  
a button comprising a particular word, a text, a symbol or graphics  
and a document) in a broad sense.

20 Specifically, the hypertext is a collection of linked texts  
(electronified documents) and is the basic concepts for  
multimedia software with which texts hierarchically related to  
each other can simultaneously be referred from multiple sites  
to in an arbitrary order through a link structure. Therefore,  
25 each document of this hypertext includes information for forming

the link structure (namely, information as a hypertext) in addition to information visually provided to users.

By the way, in the field of computer technology, a search engine is provided as a means of searching for a document linked  
5 to a particular word (keyword) from vast number of undefined documents. It should be noted that in the conventional type of search engine a keyword is not entered in a document but it is entered in a particular dialog box.

However, although there are many advantages in the  
10 hypertext, a jump can be made only to a linked page (document), and so, when a user wants to have information related to any word not linked to some other documents, it is necessary to actuate a search engine or the like and perform a search (a keyword search or a global search), which makes the operation complicated, and  
15 also linkage among documents is not easy therefore it is not always easy to use the hypertext.

In the conventional type of hypertext, an operation with one button can be linked only to one document (page), which does not allow a plurality of documents to be referred to with one  
20 button. Especially, the Intranet documents are highly specialized so that there may exist a plurality of information that the user wants to refer to with one button, hence convenience of use and workability are not sufficient for the situation.

In the conventional type of search engine, a user interface  
25 is of inferior performance as compared to a hypertext and the

operating method is also different, and so, insufficient workability and convenience are found in actuation of a search engine and performance of a search during operation of the hypertext. Especially, in a browser function for displaying and  
5 operating a hypertext on a screen, it is general to operate only with a pointing device such as a mouse, so that entry of a keyword into a dialog box of the search engine through a keyboard disturbs the flow of work.

#### 10 SUMMARY OF THE INVENTION

It is an object of the present invention to get, for the purpose of solving the problems, enhanced convenience in document information management as well as in a document search, improved operability of a keyword search, and further promoted workability  
15 when a search is performed through a hypertext by performing a search (a keyword search and/or a global search) with any word which is not linked to other documents as a keyword through a clicking operation thereon and enabling jumping from the clicked word to a page as an object to be searched like that with a button  
20 in a hypertext.

In the document information management system according to the present invention, a search-engine-compatible interface unit makes a word in a document displayed on the screen to be specified, transfers the specified word to a search engine as  
25 a keyword to be used in the search engine, receives a search result

from the search engine, and displays the search result on the screen. On the other hand, a browser-compatible interface unit performs a search by using the keyword transferred from a browser and transfers a search result to the browser. Thus, by performing  
5 a search (a keyword search and/or a global search) with a word which is not linked to any document as a keyword through a clicking operation on the word, and enabling jumping from the clicked word to a page as an object to be searched like that with a button of a hypertext, it is possible to improve convenience in document  
10 information management as well as in a document search.

In the document information management system according to another aspect of the present invention, a search-engine-compatible interface unit makes a word in a document displayed on the screen to be specified, actuates a search engine and make  
15 the engine perform a search with the specified word as a keyword to be used in the search engine, receives a search result from the search engine and displays the result on the screen. Thus, by performing a search with a word which is not linked to any document as a keyword through a clicking operation on the word,  
20 and enabling jumping from the clicked word to a page as an object to be searched like that with a button of a hypertext, it is possible to improve convenience in document information management as well as in a document search.

Other objects and features of this invention will become  
25 apparent from the following description with reference to the

accompanying drawings.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a view showing a document information management  
5 system according to Embodiment 1 of the present invention  
represented in the claim;

Fig. 2 is an explanatory view showing a concrete example  
of hardware configuration of the document information management  
system according to Embodiment 1;

10 Fig. 3 is a schematic flow chart of the document information  
management system according to Embodiment 1;

Fig. 4 is an explanatory view showing a display example  
of a pop-up menu according to Embodiment 1;

Fig. 5 is a view showing a document information management  
15 system according to Embodiment 2 of the present invention  
represented in the claim;

Fig. 6 is a schematic flow chart of the document information  
management system according to Embodiment 2;

Fig. 7 is a schematic flow chart of a document information  
20 management system according to Embodiment 3;

Fig. 8 is an explanatory view showing a concrete example  
of processing in steps S702 to S705 of Fig. 7;

Fig. 9 is an explanatory view showing a display example  
of a keyword according to Embodiment 3;

25 Fig. 10 is a schematic flow chart of a document information

management system according to Embodiment 4;

Fig. 11A and Fig. 11B are explanatory views showing how a word is identified by a search-engine-compatible interface unit according to Embodiment 5;

5 Fig. 12 is a schematic flow chart of a document information management system according to Embodiment 6;

Fig. 13 is a schematic flow chart of a document information management system according to Embodiment 7;

Fig. 14 is an explanatory view showing an example of how  
10 to specify (identify) search criteria and output criteria through a mouse operation in Embodiment 7;

Fig. 15 is an explanatory view showing an example of how to specify (identify) search criteria and output criteria through a mouse operation in Embodiment 7;

15 Fig. 16 is an explanatory view showing an example of how to specify (identify) search criteria and output criteria through a mouse operation in Embodiment 7;

Fig. 17 is an explanatory view showing an example of how to specify (identify) search criteria and output criteria through  
20 a mouse operation in Embodiment 7; and

Fig. 18 is an explanatory view showing an example of how to specify (identify) search criteria and output criteria through a mouse operation in Embodiment 7.

## DESCRIPTION OF THE PREFERRED EMBODIMENTS

Detailed description is made hereinafter for the document information management system according to the present invention in the order of Embodiment 1 to Embodiment 7 with reference to  
5 the accompanied drawings.

The document information management system according to Embodiment 1 is a system for getting enhanced convenience in document information management as well as in a document search, improved operability of a search, and further promoted  
10 workability when a search is performed through a hypertext by using a document displayed on the screen of a browser as a front end of a search engine and displaying a search result by the search engine on the screen of the browser.

Fig. 1 shows the document information management system  
15 according to Embodiment 1 represented in the claim which comprises at least one database or a filing unit (described only as the database hereinafter to make description simpler) 101 for storing therein a plurality of document files comprising a hypertext document written in hypertext language such as HTML  
20 and/or a non-hypertext document prepared with an application such as word processor, a browser 102 having a browsing function for displaying and operating the hypertext document and the non-hypertext document on a screen, and a search engine 103 having a search-engine function of searching for a corresponding  
25 document through the database (or the filing unit) 101 according

to a specified keyword.

It should be noted that the browser 102 has a search-engine-compatible interface unit 104 for specifying a word in a document displayed on the screen, transferring the specified  
5 word to the search engine 103 as a keyword to be used in the search engine 103, receiving a search result from the search engine 103, and displaying the search result on the screen. Herein, the browser 102 receives a list for a pop-up menu from the search engine 103 (browser-compatible interface unit 105) as a search  
10 result by the search (keyword search or/and global search), and displays the pop-up menu in a space adjacent to the specified word.

The search engine 103 has a browser-compatible interface unit 105 for performing a search by using the keyword transferred  
15 from the browser 102 (to be precise search-engine-compatible interface unit 104) and transferring a search result to the browser 102 (to be precise search-engine-compatible interface unit 104). Herein, the search engine 103 prepares a list for a pop-up menu and transfers the list to the search-engine-  
20 compatible interface unit 104.

Fig. 2 shows a concrete example of hardware configuration of the document information management system according to Embodiment 1, in which a plurality of Web servers 201 connected to a network such as the Intranet can be used as databases 101.  
25 This Web server 201 has a large-capacity hard disk device for



storing the documents (hypertext documents and non-hypertext documents) to be used in the system.

As the browser 102, a terminal unit 202 such as a personal computer with built-in application software for realizing a browser function can be used. This terminal unit 202 has at least a display for displaying a document thereon, a pointing device such as a mouse for operating the screen, and a communication unit for carrying out communications through a network Net. Of course, the search-engine-compatible interface unit 104 is incorporated in the terminal unit 202 as one function of application software for realizing a browser function. A number of terminal units 202 is not particularly specified, and so it is needless to say that a plurality of terminal units 202 can be provided on the network Net as shown in the figure.

As the search engine 103, a terminal unit 203 such as a personal computer with built-in application software for realizing a search engine function can be used. This terminal unit 203 is connected to the network Net, and a plurality of units may be provided. One terminal unit can also be shared for the terminal unit 202 as the browser 102 and the terminal unit 203 as the search engine 103. More specifically, the application software for realizing the search engine function may be incorporated in the terminal unit 202 with the built-in application software for realizing the browser function. Of course, the browser-compatible interface unit 105 is

incorporated in the terminal unit 203 (together with the search engine 103) as one function of application software for realizing a browser function.

Description is made for the operation based on the configuration described above with reference to Fig. 3. Fig. 3 shows a schematic flow chart of the document information management system according to Embodiment 1. In Embodiment 1, when a keyword is inputted through the browser 102, the browser-compatible interface unit 105 actuates the search engine 103, and a search with the keyword is started.

At first, the browser 102 displays a document on the screen (S301). The search-engine-compatible interface unit 104 starts its operation in this state.

When a word is identified (specified) through a mouse operation of the terminal unit 202 (S302), the search-engine-compatible interface unit 104 recognizes the word as a keyword (S303), transfers the keyword to the search engine 103 (S304), and displays the recognized keyword (S305). Identification (specification) of a word herein is performed, for example, by clicking the second mouse button in order to differentiate the above processing from processing of the hypertext document. The keyword is provided in a reverse video or is underlined so that it can be discriminated from other words.

On the other hand, when the browser-compatible interface unit 105 of the search engine 103 receives a keyword from the

search-engine-compatible interface unit 104 of the browser 102  
it actuates the search engine 103 and sets the keyword in a  
keyword-entry item of the search engine 103 (S351). Herein, the  
browser-compatible interface unit 105, although it is a part of  
5 the application for the search engine 103, always stays resident  
on the terminal unit 203 and also plays a role as a trigger to  
actuate the main system of the search engine 103 when the keyword  
is transferred from the search engine 103.

When the keyword is set in step S351, the search engine  
10 103 executes a search operation with the set keyword (S352). It  
should be noted that the search with the keyword in this case  
is performed based on the assumption that an object to be searched  
by the search engine 103 (the database 101 to be searched) is  
previously set and specified. Therefore, a search can be  
15 performed across all the databases 101 connected to the network  
Net as objects to be searched through setting or specification  
thereof, and can also be performed across only a particular  
database 101 as an object to be searched. Setting and  
specification of the database 101 as an object to be searched  
20 may be carried out with the browser 102 although it is not  
mentioned herein.

Then, the browser-compatible interface unit 105 prepares  
a list for a pop-up menu from a search result with the keyword  
(S353), and transfers the prepared list to the browser 102 (S354).  
25 It should be noted that document names having been searched and

URLs (addresses) for those documents are described on the list.

When the search-engine-compatible interface unit 104 of the browser 102 receives a list from the browser-compatible interface unit 105 it develops the list to a hypertext and overwrites (displays) the list in a space adjacent to the keyword of the original document as a pop-up menu (S306). It should be noted that Fig. 4 shows a display example of a pop-up menu 402 appearing on the screen 401 of the browser 102.

When the user identifies a desired document out of the pop-up menu as indicated by the cursor 403 in Fig. 4 through the mouse operation, the search-engine-compatible interface unit 104 notifies the URL (identification of the document) of the corresponding document described on the list to the browser 102 (S307). The browser 102 accesses the database 101 (Web server 201) with a document in the linked address stored therein according to the URL, and requests the database 101 to transfer the source code of the corresponding document (S308).

The database 101 having received the request from the browser 102 to transfer the source code transfers the source code of the corresponding document to the browser 102 (S391).

The browser 102 after receiving the source code develops the code to a hypertext document or a non-hypertext document (application document), and displays either of the documents on the screen (S309).

In accordance with Embodiment 1 as described above, a

search is performed by using a word which is not linked to other documents as a keyword by clicking on the word, and it is possible to jump from the clicked word to a page as an object to be searched, which is just like the operation with a button of a hypertext.

5 Therefore, it is possible to enhance convenience in document information management as well as in a document search, improve operability of a keyword search, and further promote workability when a search is performed through a hypertext. In other words, it is possible to use a document displayed on the screen of a  
10 browser 102 as a front end of the search engine 103 and display a search result by the search engine 103 on the screen of the browser 102.

It should be noted that the browser-compatible interface unit 105 can be realized by rewriting API (Application  
15 Programming Interface) of the existing search engine 103. Also, a search-engine-compatible interface unit 104 of the browser 102 can be realized by being added to the existing browser as plug-in software.

The document information management system according to  
20 Embodiment 2 is an example having the same effect as that of Embodiment 1 in which the existing search engine 103 not having a browser-compatible interface unit 105 is used. Namely, in the document information management system according to Embodiment 1, the browser 102 has the search-engine-compatible interface  
25 unit 104 and the search engine 103 has the browser-compatible

interface unit 105, through which the document appearing on the screen of the browser 102 is usable as a front end of the search engine 103, but in Embodiment 2, by expanding the function of the search-engine-compatible interface unit 104 of the browser  
5 103, the existing search engine 103 without a browser-compatible interface unit 105 incorporated therein can be used.

It should be noted that an example of basic hardware configuration is the same as that in Embodiment 1, so that description is made herein only for the different sections.

10 Fig. 5 shows a the document information management system according to Embodiment 2 represented in the claim which comprises at least one database or a filing unit (described only the database hereinafter to make description simpler) 101 for storing therein a plurality of document files comprising a  
15 hypertext document written in a hypertext language such as HTML and/or a non-hypertext document prepared with an application such as word processor, a browser 102 having a browsing function for displaying and operating the hypertext document and non-hypertext document on a screen, and a search engine 103 having  
20 a search-engine function of searching for a corresponding document through the database 101 according to a specified keyword.

It should be noted that the browser 102 has a search-engine-compatible interface unit 106 for specifying a word in  
25 a document displayed on the screen, actuating the search engine

103 to make the engine perform a search with the specified word as a keyword to be used in the search engine 103, receiving a result of the search from the search engine 103, and displaying the search result on the screen.

5           Description is made for the operation based on the configuration described above with reference to Fig. 6. Fig. 6 shows a schematic flow chart of the document information management system according to Embodiment 2. In Embodiment 2, when receiving a keyword, the browser 102 actuates the search  
10 engine 103, and a search with the keyword is started.

At first, the browser 102 displays a document on the screen (S601). The search-engine-compatible interface unit 106 starts its operation in this state.

When a word is identified (specified) through a mouse  
15 operation of the terminal unit 202 (S602), the search-engine-compatible interface unit 106 recognizes the word as a keyword (S603), and actuates the search engine 103 (S604). Herein, the actuated search engine 103 transfers a screen for entering a keyword to the search-engine-compatible interface  
20 unit 106 as processing executable on the system (S651).

Even if the search-engine-compatible interface unit 106 receives the keyword entering screen from the search engine 103 it does not display this screen (namely, the processing of displaying the screen is skipped), but transfers the keyword  
25 recognized in step S603 to the search engine 103 (S605), and

displays the recognized keyword on the screen (S606). Identification (specification) of a word herein is performed, for example, by clicking the second mouse button in order to differentiate the above processing from processing of the  
5   hypertext document. The keyword is provided in a reverse video or is underlined so that it can be discriminated from other words.

On the other hand, when the search engine 103 receives a keyword from the search-engine-compatible interface unit 106 it performs the processing of a search with the received keyword  
10   (S652). It should be noted that the search with the keyword in this case is performed based on the assumption that an object to be searched by the search engine 103 (the database 101) is previously set and specified. Therefore, a search can be performed across all the databases 101 connected to the network  
15   Net as objects to be searched through setting or specification thereof, and can also be performed across only a particular database 101 as an object to be searched. Setting and specification of the database 101 as an object to be searched may be carried out with the browser 102 although it is not  
20   mentioned herein.

Then the search engine 103 transfers a result of the search to the search-engine-compatible interface unit 106 (S653).

While the search-engine-compatible interface unit 106 cuts out, when receiving the keyword-search result, the button from  
25   the search result to prepare a list for a pop-up menu (S607),



develops the list to a hypertext, and overwrites (displays) the list in a space adjacent to the keyword of the original document as a pop-up menu (S608).

When the user identifies a desired document out of the  
5 pop-up menu through the mouse operation, the search-engine-compatible interface unit 106 notifies the URL (identification of the document) of the corresponding document described on the list to the browser 102 (S609). The browser 102 accesses the database 101 (Web server 201) with a document in the linked address  
10 stored therein according to the URL, and requests the database 101 to transfer the source code of the corresponding document (S610).

The database 101 having received the request from the browser 102 to transfer the source code transfers the source  
15 code of the corresponding document to the browser 102 (S691).

When the browser 102 receives the source code it develops the code to a hypertext document or a non-hypertext document (application document), and displays either of the documents on the screen (S611).

20 In accordance with Embodiment 2 as described above, in addition to the same effect as that in Embodiment 1, the existing search engine 103 can be used as it is, which makes it easy to apply the invention to the existing system, and the convenience of use can further be improved.

25 In the document information management system according

to Embodiment 3, a document file for a document displayed on the screen and used for specifying a keyword by the browser has a word table prepared by previously cutting out words included in the document file. The search-engine-compatible interface unit  
5 selects in the document displayed on the screen a plurality of expanded character strings each obtained by adding characters or character strings before and behind a character pointed by a cursor thereto, compares the selected strings to the words in the word table added to the document to evaluate the strings,  
10 and identifies an optimal word which can be assumed as the one being pointed by the cursor as a keyword to be used in the search engine.

The document information management system according to the present invention requires a function of identifying a word  
15 specified in the browser (search-engine-compatible interface unit), but this function of identifying a word can be realized in various methods, so that, a method is not particularly defined in Embodiment 1 and Embodiment 2. Therefore, in Embodiment 3, the function of identifying a word preferable to the document  
20 information management system is explained concretely.

It should be noted that the basic configuration and the operation thereof are the same as those in Embodiment 1, so that detailed description is made herein only for the different sections.

25 By the way, it is not easy to identify a word in a document

such as the one written in Japanese. In such a case the method of identifying a word includes a method of preparing a particular document structure previously produced according to a use and a method of using an existing document structure. Herein, the document structure for the browser 102 of the document information management system according to the present invention makes it easy to identify a word by using a document (a hypertext document or a non-hypertext document) having a word table prepared by previously cutting out words included in a document file.

Added to a document file handled in the document information management system according to Embodiment 3 is a word table prepared by previously cutting out words in a document file. When a global search is to be performed across any document in Japanese, Chinese, or Korean language as an object to be searched, an index file such as a character element table and a word table is added to the document to be searched. Therefore, all of the documents handled in the system has a word table added thereto respectively, and this word table functions as a word table used for identifying a word in the document on the screen of the browser 102 and functions as an index file used for a global search in the document to be searched.

The preparation of the word table will be carried out as follows with terminal units 202, 203 and a Web server 201 connected to the network Net.

At first, words are cut out of a document registered anew by using a technique such as morpheme analysis at the point of time when a prepared document is registered or during the time zone such as night time when the document information management system is not used, then, a word table is prepared with the cut-out words, and the word table is added to the original document file, and the rewritten data is registered.

The processing as described above is always carried out, which allows word tables to be added to all the document files existing in the document information management system according to Embodiment 3.

Description is made for the operation based on the configuration described above with reference to Fig. 7. Fig. 7 shows a schematic flow chart of the document information management system according to Embodiment 3. In Embodiment 3, when a keyword is identified by the browser 102, the browser-compatible interface unit 105 actuates the search engine 103, and a search with the keyword is started.

At first, the browser 102 displays a document on the screen (S701). The search-engine-compatible interface unit 104 starts its operation in this state.

When a character is identified (specified) through a mouse operation of the terminal unit 202 (S702), the search-engine-compatible interface unit 104 cuts out a character string obtained by adding some characters before and after the character

thereto (adjacent character string) (S703), compares the cut-out character string to the word table (S704), cuts out the character string coincident with any word in the word table as a word, and identifies the word (S705).

5           Herein, a concrete example of processing in steps S702 to S705 is shown with reference to Fig. 8. As shown in Fig. 8, for example, a cursor 801 is moved to an arbitrary character (herein a character "A" in the word "MARKED") in the document displayed on the screen, a mouse button is clicked, and the character  
10 (character "A") at the position of the cursor is specified. The character at the position of the cursor specified as described above is identified as an identified character 702A in step S702 (Refer to Fig. 7). It should be noted that identification of a word herein is performed, for example, by clicking the second  
15 mouse button in order to differentiate the above processing from processing of the hypertext document.

          Then, a character string obtained by adding some characters before and after the identified character thereto (adjacent character string) is cut out. It is conceivable that there are  
20 various methods as the method of cutting out a character, but there is the simplest one in which punctuation marks before and after an identified character are decided as identifiers and a character string between the punctuation marks can be cut out as an adjacent character string. In this case, the adjacent  
25 character string cut-out in step S703 is as follows: "when a

plurality of buttons are marked a plurality of pages are outputted onto the display or to the printer." as indicated by the reference numeral 703A in Fig. 7.

Then, the adjacent character string is compared with the words in the word table, and the character string coincident with any word in the word table is cut out as a word, and the word is identified. This identification of a word can be carried out as follows.

1) A character string "MA" obtained by adding one character before the specified character "A" is decided as a temporary word, and comparison is made as to whether there is any word coinciding with this temporary word in the word table or not, and when it is determined that there is a coinciding word in the word table, the temporary word is identified as a word.

2) When it is determined that there is no coinciding word in the word table, a character string "AR" obtained by adding one character behind the specified character "A" is decided as a temporary word, and comparison is made as to whether there is any word coinciding with this temporary word in the word table or not, and when it is determined that there is the coinciding word, the temporary word is identified as a word.

3) When it is determined that there is no coinciding word therein, a character string "MAR" obtained by adding one character before and after the specified character "A" is decided as a temporary word, and comparison is made as to whether there

is any word coinciding with this temporary word in the word table or not, and when it is determined that there is the coincident word therein, the temporary word is identified as a word.

Thereafter and on, characters are added in the same manner  
5 as described above, and the processing is repeated until the cut-out character string is coincident with any word in the word table or until 'the cut-out character string' becomes the same as 'the adjacent character string'.

Herein, a word "MARK" is identified as indicated by the  
10 reference numeral 705A in Fig. 7 during the identification of a word in step S705.

In English language the words in a sentence are separated by a space, so that a character string sandwiched between spaces can without difficulty be considered as a word. However, in  
15 languages like Japanese, Chinese or Korean, words in a sentence are not separated by space as in English. Therefore, in the documents written in these languages, it is difficult to decide where a word starts and where it ends in a sentence. In the above described method, a character is added before or behind a  
20 specified character, so that, a stage is reached when a word having a meaning is formed and the word coincides with a word in the word table. Though an English language document is shown in Fig. 8 for simplification in explanation, this method is more effective for extracting a word from the documents written in  
25 Japanese, Chinese or Korean languages.

The identified word "MARK" is recognized as a keyword in step S706 of Fig. 7, and the keyword is transferred to the search engine 103. The recognized keyword is also displayed on the screen in step S707. Fig. 9 shows an example of display of a keyword. The keyword is provided in a reverse video or is underlined so that it can be distinguished from other words.

On the other hand, when the browser-compatible interface unit 105 of the search engine 103 receives the keyword from the search-engine-compatible interface unit 104 of the browser 102 it actuates the search engine 103 and sets the keyword in a keyword-entry item of the search engine 103 (S751). Herein, the browser-compatible interface unit 105, although it is a part of the application for the search engine 103, always stays resident on the terminal unit 203 and also plays a role as a trigger to actuate the main system of the search engine 103 when the keyword is transferred from the search engine 103.

When the keyword is set in step S751, the search engine 103 executes a search operation with the set keyword (S752).

Then, the browser-compatible interface unit 105 prepares a list for a pop-up menu from a search result with the keyword (S753), and transfers the prepared list to the browser 102 (S754). It should be noted that document names having been searched and URLs (addresses) for those documents are described in this list.

When the search-engine-compatible interface unit 104 of the browser 102 receives a list from the browser-compatible



interface unit 105 it develops the list to a hypertext and overwrites (displays) the list in a space adjacent to the keyword of the original document as a pop-up menu (S708).

When the user identifies a desired document out of the  
5 pop-up menu through the mouse operation, the search-engine-compatible interface unit 104 notifies the URL (identification of the document) of the corresponding document described in the list to the browser 102 (S709). The browser 102 accesses the database 101 (Web server 201) with a document in the linked address  
10 stored therein according to the URL, and requests the database 101 to transfer the source code of the corresponding document (S710).

The database 101 having received the request from the browser 102 to transfer the source code transfers the source  
15 code of the corresponding document to the browser 102 (S791).

When the browser 102 receives the source code it develops the code to a hypertext document or a non-hypertext document (application document), and displays either of the documents on the screen (S711).

20 In accordance with Embodiment 3 as described above, in addition to the same effect as that in Embodiment 1, a document file has a word table prepared by previously cutting out words included in the document file, and the search-engine-compatible interface unit selects, in the document displayed on the screen,  
25 a plurality of expanded character strings each obtained by adding

characters or character strings before and after a character pointed by a cursor thereto, compares the selected strings to the words in the word table to evaluate the strings, and identifies an optimal word which can be assumed as the one being pointed  
5 by the cursor as a keyword to be used in the search engine, therefore a word as a keyword can accurately and efficiently be identified.

The document information management system according to Embodiment 4 concretely shows another example of a function for  
10 identifying a word preferable to the document information management system according to the present invention. It should be noted that the basic configuration and the operation thereof are the same as those in Embodiment 1, so that detailed description is made herein only for the different sections.

15 In the document information management system according to Embodiment 4, a hypertext document handled in the system has a dummy button obtained by previously cutting out a word included in the document and setting a target for linking the word thereto as particular address information for responding to actuation  
20 of the search engine provided therein. Further, the search-engine-compatible interface unit identifies, when the dummy button is pointed by the cursor, a word corresponding to the dummy button as a keyword to be used in the search engine, and also actuates the search engine according to the URL for the dummy  
25 button.

Thus, the hypertext document handled in Embodiment 4 has a previously set dummy button. It should be noted that any word having been set as an ordinary button in a hypertext document is not set as a dummy button, but all the main words other than the button therein are set as dummy buttons. A linked URL of this dummy button is set in the search engine 103 or in a common particular URL.

The setting and preparation of this dummy button is carried out as follows with terminal units 202, 203 and the Web server 201 connected to a network Net. At first, a word is cut out of a document registered anew by using a technique such as morpheme analysis at the point of time when a prepared document is registered or during the time zone such as night time when the document information management system is not used, then, the cut-out word is identified as a dummy button, the linked URL thereof is set, the dummy button is added to the original document file, and the rewritten data is registered.

The processing as described above is always carried out, which allows dummy buttons to be added to all the hypertext documents existing in the document information management system according to Embodiment 4.

Description is made for the operation based on the configuration described above with reference to Fig. 10. Fig. 10 shows a schematic flow chart of the document information management system according to Embodiment 4. In Embodiment 4,

when a dummy button is identified by the browser 102, the browser-compatible interface unit 105 actuates the search engine 103, and a search with the keyword is started.

At first, the browser 102 displays a document on the screen (S1001). The search-engine-compatible interface unit 104 starts its operation in this state.

When a dummy button is identified (specified) through a mouse operation of the terminal unit 202 (S1002), the search-engine-compatible interface unit 104 accesses to the linked URL set in the dummy button (S1003), transfers the character string (word) for the dummy button to the linked location as a keyword (S1004), and displays the identified dummy button (keyword) on the screen (S1005). It should be noted that a keyword is provided in a reverse video or is underlined so that it can be distinguished from the button and other words in a hypertext.

On the other hand, when the access is made from the search-engine-compatible interface unit 104 of the browser 102 in step S1003, the browser-compatible interface unit 105 of the search engine 103 actuates the search engine 103 (S1051), recognizes the keyword when it receives the keyword from the search-engine-compatible interface unit 104 in step S1004 and sets the keyword in a keyword-entry item of the search engine 103 (S1052). Herein, the browser-compatible interface unit 105, although it is a part of the application for the search engine

103, always stays resident on the terminal unit 203 and also plays a role as a trigger to actuate the main system of the search engine 103 when the keyword is transferred from the search engine 103.

When the keyword is set in step S1052, the search engine  
5 103 executes a search operation with the set keyword (S1053).

Then, the browser-compatible interface unit 105 prepares a list for a pop-up menu from a search result with the keyword (S1054), and transfers the prepared list to the browser 102 (S1055). It should be noted that document names having been  
10 searched and URLs (addresses) for those documents are described in this list.

When the search-engine-compatible interface unit 104 of the browser 102 receives a list from the browser-compatible interface unit 105 it develops the list to a hypertext and  
15 overwrites (displays) the list in a space adjacent to the keyword of the original document as a pop-up menu (S1006).

When the user identifies a desired document out of the pop-up menu through the mouse operation, the search-engine-compatible interface unit 104 notifies the URL (identification  
20 of the document) of the corresponding document described in the list to the browser 102 (S1007). The browser 102 accesses the database 101 (Web server 201) with a document in the linked address stored therein according to the URL, and requests the database to transfer the source code of the corresponding document  
25 (S1008).

The database 101 having received the request from the browser 102 to transfer the source code transfers the source code of the corresponding document to the browser 102 (S1091).

When the browser 102 receives the source code it develops  
5 the code to a hypertext document or a non-hypertext document (application document), and displays either of the documents on the screen (S1009).

In accordance with Embodiment 4 as described above, in addition to the same effect as that in Embodiment 1, a hypertext  
10 document has a dummy button obtained by previously cutting out a word included in the document and setting the linked address of the word as particular address information (URL) for responding to actuation of the search engine, therefore a word as a keyword can efficiently be identified and also a search engine  
15 can easily be actuated.

Embodiment 5 shows another example of a function for identifying a word preferable to the document information management system according to the present invention, and a search engine-compatible interface unit cuts out, in a document  
20 displayed on the screen, a character string from a header character to an end character indicated by the cursor as a word, and identifies the cut-out word as a keyword to be used in the search engine. It should be noted that the basic configuration and the operation thereof are the same as those in Embodiment  
25 1, so that description is made herein only for a method of

identifying a word by the search-engine-compatible interface unit 104.

Fig. 11A and Fig. 11B are explanatory views showing how a word is identified by the search-engine-compatible interface unit 104 according to Embodiment 5. The search-engine-compatible interface unit 104 detects, as shown in Figs. 11A and 11B, a first position 1101 and a second position 1102 of the cursor in the document on the screen of the browser 102, cuts out a character string from the character at the first position 1101 to the character at the second position 1102 as a specified word, and sets the word as an identified word 1103.

More specifically, a method of clicking ON a mouse button at the first position 1101 and clicking OFF the mouse button at the second position 1102 and a method of clicking the mouse button at the first position 1101 and again clicking the mouse button at the second position 1102 may be used. Also identification of a word is performed, for example, by clicking the second mouse button in order to differentiate the above processing from processing of the hypertext document. Further, the identified word is provided in a reverse video or is underlined.

In accordance with Embodiment 5 as described above, there is no need to have a particular document structure, therefore an existing hypertext document or a non-hypertext document (application document) can be used as it is. In addition, a word as a keyword can efficiently be identified.

In the document information management system according to Embodiment 6, the search-engine-compatible interface unit cuts out, in the document displayed on the screen, an optimal word through evaluation of an expanded character string obtained  
5 by adding characters or character strings before and behind a character pointed by a cursor thereto by using a word cut-out technique such as morpheme analysis, and identifies the cut-out word as a keyword to be used in the search engine.

It should be noted that the basic configuration and the  
10 operation thereof are the same as the document information management system according to Embodiment 3, so that description is made herein only for the different sections.

Fig. 12 shows a schematic flow chart of the document information management system according to Embodiment 6, in which  
15 step S1201 is provided in place of step S704 of the schematic flow chart in Embodiment 3 shown in Fig. 7, and the other steps are the same as those in Embodiment 3.

When a character is identified (specified) through a mouse operation of the terminal unit 202 (S702), the search-  
20 engine-compatible interface unit 104 cuts out a character string obtained by adding some characters before and after the character thereto (adjacent character string) (S703), cuts out a word through morpheme analysis of the cut-out character string for evaluation (S1201), and identifies an optimal word (S705).

25 Herein, the adjacent character string cut-out in step S703



is as follows: "when a plurality of buttons are marked, a plurality of pages are outputted onto the display or to the printer." as indicated by the reference numeral 703A in Fig. 12. When this adjacent character string is subjected to the morpheme analysis, the word "MARK" can be identified as indicated by the reference numeral 705A in Fig. 12.

In accordance with Embodiment 5 as described above, the search-engine-compatible interface unit 104 cuts out an optimal word through evaluation of a character string by using a word cut-out technique such as morpheme analysis, and identifies the cut-out word as a keyword to be used in the search engine, therefore a word as a keyword can accurately and efficiently be identified.

The document information management system according to Embodiment 7 enables entry of search criteria and output criteria much more as compared to those in Embodiment 1 to Embodiment 6. It should be noted that the basic configuration is the same as that of Embodiment 1, so that description is made herein only for the different sections.

In the document information management system according to Embodiment 7, the search-engine-compatible interface unit 104 displays a menu for specifying search criteria and output criteria for the search engine 103 in a space adjacent to the specified word, and performs a search and output according to the specified word as well as to the menu items specified by the

cursor.

Description is made for the operation based on the configuration described above with reference to Fig. 13. Fig. 13 shows a schematic flow chart of the document information management system according to Embodiment 7. In Embodiment 7, when a keyword is identified by the browser 102, the browser-compatible interface unit 105 actuates the search engine 103, and a search with the keyword is started.

At first, the browser 102 displays a document on the screen (S1301). The search-engine-compatible interface unit 104 starts its operation in this state.

When a word is identified (specified) through a mouse operation of the terminal unit 202 (S1302), the search-engine-compatible interface unit 104 recognizes, the word as a keyword (S1303), and displays the recognized keyword on the screen (S1304).

Then, a pop-up menu for specifying search criteria as well as output criteria is overwritten and displayed adjacent to the recognized keyword (S1305). More specifically, as shown in Fig. 14, when the keyword 1401 (for example, "LINKED") is displayed, the pop-up menu for the search criteria as well as output criteria is displayed adjacent to the corresponding keyword. By clicking on a desired character on this pop-up menu (not shown), the corresponding item can be selected. In the search criteria, it is possible to specify a category and a type of DB (database to

be searched), and in the output criteria, it is possible to specify output of related information and data for the related information and also specify a target to be outputted of data (a device to which the data is outputted).

5           When the search criteria and output criteria are identified through the mouse operation described later in step S1306, the search-engine-compatible interface unit 104 transfers the recognized keyword and the identified search criteria to the search engine 103 (S1307).

10           On the other hand, when the browser-compatible interface unit 105 of the search engine 103 receives a keyword from the search-engine-compatible interface unit 104 of the browser 102 it actuates the search engine 103 and sets the keyword for the search engine 103 and the search criteria as well as output  
15   criteria therein (S1351). Herein, the browser-compatible interface unit 105, although it is a part of the application for the search engine 103, always stays resident on the terminal unit 203 and also plays a role as a trigger to actuate the main system of the search engine 103 when the keyword is transferred from  
20   the search engine 103.

          When the keyword and the search criteria as well as output criteria are set in step S1351, the search engine 103 executes a search operation with the set keyword and search criteria as well as output criteria (S1352).

25           Then, the browser-compatible interface unit 105 determines

whether output of data for related information is specified as an output criterion or not (S1353). Herein, when it is determined that the output of data for related information is set, the browser-compatible interface unit 105 prepares a list for a pop-up menu from a keyword-search result (S1355), and transfers the prepared list to the browser 102 (S1356). When it is determined that the output of data for related information is not set, the browser-compatible interface unit 105 transfers only the URL for the related information to the browser 102 (S1354).

When the search-engine-compatible interface unit 104 of the browser 102 receives a list from the browser-compatible interface unit 105 it develops the list to a hypertext and overwrites (displays) the list in a space adjacent to the keyword of the original document as a pop-up menu (S1308).

When the user identifies a desired document out of the pop-up menu through the mouse operation, the search-engine-compatible interface unit 104 notifies the URL (identification of the document) of the corresponding document described in the list to the browser 102 (S1309). The browser 102 accesses the database 101 (Web server 201) with a document in the linked address stored therein according to the URL, and requests the database 101 to transfer the source code of the corresponding document (S1310). Similarly, when the browser 102 receives the URL for the related information from the browser-compatible interface unit 105 it accesses the database 101 (Web server 201) with a

document in the linked address stored therein according to the URL, and requests the database 101 to transfer the source code of the corresponding document.

The database 101 having received the request from the browser 102 to transfer the source code transfers the source code of the corresponding document to the browser 102 (S1391).

When the browser 102 receives the source code it develops the code to a hypertext document or a non-hypertext document (application document), and displays either of the documents on the screen (S1311).

Next, description is made more specifically for a method of specifying (identifying) search criteria as well as output criteria through a mouse operation with reference to Fig. 14 to Fig. 18.

Specification of the search criteria as well as output criteria is performed, for example, by clicking the second mouse button in order to differentiate the above processing from processing of the hypertext document. The specified word is provided in a reverse video or is underlined.

As the search criteria, a simple search equation can be used. For example, OR search does not require a search equation while an AND search is performed through the operation described below. At first, it is assumed, as shown in Fig. 14, that the criteria indicated by the pop-up menu 1402 are set for the keyword 1401 and the criteria (default) indicated by the pop-up menu

1404 are set for the keyword 1403. In this status, the cursor is pointed to the keyword 1403 having been specified and the second mouse button is clicked ON, and the cursor is moved to the keyword 1405 which is not specified yet and the second mouse button is  
5 clicked OFF, and then an AND sign " $\cap$ " is generated over the keyword 1403 to the keyword 1405. At this point of time, the same contents as those of the pop-up menu 1404 are displayed as a pop-up menu 1406. It should be noted that only the pop-up menu 1404 may be displayed for the AND search.

10 As some other method, the cursor is pointed to the keyword 1403 having been specified and the second mouse button is clicked ON, and the cursor is moved to the keyword 1405 which is not specified yet and the second mouse button is clicked OFF, and then the keyword 1403 is moved to the position of the keyword  
15 1405 and automatically returned to the original position as soon as the second mouse button is clicked OFF (drag & drop), and the AND sign " $\cap$ " is generated over the keyword 1403 to the keyword 1405. At this point of time, the same contents as those of the pop-up menu 1404 are displayed as the pop-up menu 1406.

20 Some other search criteria as well as output criteria are also specified by clicking thereon. For example, as shown in Fig. 15, "Related information, DB2, Category C, Display" is specified in the pop-up menu 1404 of the AND search for the keyword 1403 and keyword 1405. This specification indicates "AND of the  
25 two keywords is searched in the category C of the DB2, and the

data for related information is outputted to the display (during its specification)". The specified search criteria as well as output criteria is displayed like the pop-up menu 1402 for the keyword 1401. It should be noted that the pop-up menu 1402  
5 indicates "One keyword is searched in the category A of the database DB1, and the related information is outputted to the printer 1 (after its specification)".

When the contents of a pop-up menu (search criteria and output criteria) having been specified once is to be corrected,  
10 the cursor is pointed to the pop-up menu having been specified and the second mouse button is clicked, and then the pop-up menu is displayed again, on which specification can be carried out again.

Then, as shown in Fig. 16, when the cursor is pointed to  
15 the "EXECUTE" button 1601 on the screen and the second mouse button is clicked, execution of a search is instructed, and the criteria therefor are identified in step S1306 of Fig. 13.

Fig. 17 shows an example of how the mouse is operated when a target to be outputted of data for related information is to  
20 be specified. As shown in steps S1355, S1356 of Fig. 13, the browser-compatible interface unit 105 of the search engine 103 transfers a search result as a list for a pop-up menu to the search-engine-compatible interface unit 104 of the browser 102. The search-engine-compatible interface unit 104 develops the  
25 list to a hypertext, and overwrites and displays the list in a

space adjacent to the original keyword as a pop-up menu 1701. The data for related information (such as a title and summary) is displayed in some space adjacent to the keyword as the pop-up menu 1701 as described above, and when the cursor is pointed to a desired data for related information on the pop-up menu 1701 and the second mouse button is clicked thereon, a pop-up menu 1801 for output criteria is displayed as shown in Fig. 18. A target to be outputted (printer 2 herein) of data can be specified by pointing the cursor to an item for the output criteria and clicking the second mouse button thereon.

It should be noted that the search criteria and output criteria shown in Fig. 14 to Fig. 18 are set as required, therefore items are not limited to the items described above. It is not always necessary to use all the items described above. Namely, only items required for each document information management system discretely constructed may be used.

In accordance with Embodiment 5 as described above, in addition to the same effect as that in Embodiment 1, the search-engine-compatible interface unit displays a menu for specifying search criteria and output criteria for the search engine in a space adjacent to a specified word, and performs a search and output according to the specified word as well as to the menu items specified by the cursor, so that workability and convenience for the user can further be improved.

As described above, in the document information management



system according to the present invention, the search-engine-compatible interface unit makes a word in a document displayed on the screen to be specified, transfers the specified word to a search engine as a keyword to be used in the search engine, receives a search result from the search engine, and displays the search result on the screen, and the browser-compatible interface unit performs a search by using the keyword transferred from a browser and transfers a search result to the browser, so that, by performing a search (a keyword search and/or a global search) with a word which is not linked to any document as a keyword through a clicking operation on the word, and enabling jumping from the clicked word to a page as an object to be searched like that with a button of a hypertext, it is possible to improve convenience in document information management as well as in a document search, operability of a keyword search, and further workability when a search is performed through a hypertext.

In the document information management system according to another aspect of the present invention, a list for a pop-up menu is prepared from a keyword-search result and the pop-up menu is displayed on the screen, so that workability and convenience for the user can further be improved.

In the document information management system according to another aspect of the present invention, a search-engine-compatible interface unit makes a word in a document displayed on the screen to be specified, actuates a search engine and make

the engine perform a search with the specified word as a keyword to be used in the search engine, receives a search result from the search engine and displays the result on the screen, so that, by performing a search with a word which is not [liked] <sup>linked</sup> to any  
5 document as a keyword through a clicking operation on the word, and enabling jumping from the clicked word to a page as an object to be searched like that with a button of a hypertext, it is possible to improve convenience in document information management as well as in a document search, operability of a  
10 keyword search, and further workability when a search is performed through a hypertext.

In the document information management system according to another aspect of the present invention, a document file has a word table prepared by previously cutting out words included  
15 in the document file, so that a word as a keyword can efficiently be identified.

In the document information management system according to another aspect of the present invention, a hypertext document has a dummy button obtained by previously cutting out a word  
20 included in the document and setting a linked target of the word as a particular address information for responding to actuation of the search engine, so that a word as a keyword can efficiently be identified and also a search engine can easily be actuated.

In the document information management system according  
25 to another aspect of the present invention, a search-engine-

compatible interface unit cuts out, in the document displayed on the screen, a character string from a header character to an end character pointed by a cursor as a word, and identifies the cut-out word as a keyword to be used in the search engine, so  
5 that a word as a keyword can efficiently be identified.

In the document information management system according to another aspect of the present invention, a search-engine-compatible interface unit evaluates, in the document displayed on the screen, an expanded character string obtained by adding  
10 characters or character strings before and after a character pointed by a cursor thereto based on a word cut-out technology such as morpheme analysis, cuts out an optimal word, and identifies the cut-out word as a keyword to be used in the search engine, so that a word as a keyword can accurately and efficiently  
15 be identified.

In the document information management system according to another aspect of the present invention, a document file has a word table prepared by previously cutting out words included in the document file, and a search-engine-compatible interface  
20 unit selects, in the document displayed on the screen, a plurality of expanded character strings each obtained by adding characters or character strings before and behind a character pointed by a cursor thereto, compares the selected strings to the words in the word table to evaluate the strings, and identifies an optimal  
25 word as if the word is pointed by the cursor as a keyword to be

used in the search engine, so that a word as a keyword can accurately and efficiently be identified.

In the document information management system according to another aspect of the present invention, a hypertext document  
5 has a dummy button obtained by previously evaluating a word included in the document based on a word cut-out technology such as morpheme analysis, cutting out the word, and setting a linked target of the word as a particular URL for responding to actuation of a search engine, and a search-engine-compatible interface unit  
10 identifies, when the dummy button is pointed by the cursor, a word corresponding to the dummy button as a keyword to be used in the search engine, and also actuates the search engine according to the URL for the dummy button, so that a word as a keyword can accurately and efficiently be identified, and also  
15 a search engine can easily be actuated.

In the document information management system according to another aspect of the present invention, a search-engine-compatible interface unit displays a menu for specifying search criteria and output criteria for a search engine in a space  
20 adjacent to a specified word, and performs a search and output according to the specified word as well as to the menu items specified by the cursor, so that workability and convenience for the user can further be improved.

In the document information management system according  
25 to another aspect of the present invention, databases or the

filing units are Web servers connected to each other through the Internet or an intranet, so that information can be collected from a wide range thereof through the Internet or the intranet.

Although the invention has been described with respect to  
5 a specific embodiment for a complete and clear disclosure, the appended claims are not to be thus limited but are to be construed as embodying all modifications and alternative constructions that may occur to one skilled in the art which fairly fall within the basic teaching herein set forth.